A test of central coherence theory: Can adults with high-functioning autism or Asperger syndrome integrate fragments of an object?

Therese Jolliffe & Simon Baron-Cohen

To cite this article: Therese Jolliffe & Simon Baron-Cohen (2001) A test of central coherence theory: Can adults with high-functioning autism or Asperger syndrome integrate fragments of an object?, Cognitive Neuropsychiatry, 6:3, 193-216, DOI: 10.1080/13546800042000124

To link to this article: http://dx.doi.org/10.1080/13546800042000124

Published online: 09 Sep 2010.

Article views: 221

View related articles

Citing articles: 29 View citing articles
A test of central coherence theory: Can adults with high-functioning autism or Asperger syndrome integrate fragments of an object?

Therese Jolliffe and Simon Baron-Cohen

University of Cambridge, Cambridge, UK

Introduction. Visuoconceptual integration was explored as a test of central coherence theory (Frith, 1989). Individuals with autism are thought to have weak central coherence so the prediction was that these individuals would show an impaired ability to integrate visual information. Method. Two groups with autistic disorder were recruited: adults with either autism or Asperger syndrome. All were normally intelligent and were matched with members of the general population of Cambridge. A modified version of the Hooper Visual Organisation Test was used in which line drawings depicting simple objects had been cut into pieces and arranged in a puzzle-like fashion. The participants were required to conceptually integrate the fragments in order to identify the object. A second condition presented just a single piece of an object and participants were required to identify objects from a single piece. Results. Both clinical groups were significantly impaired in their ability to integrate pieces holistically, but they were unimpaired in their ability to identify an object from a single piece. Conclusion. Individuals with an autistic disorder are less able to integrate visual elements. Of the two clinical groups, the autism group had the greater deficit, and it applied to the majority of the group. Possible explanations for the clinical groups’ weak central coherence are explored.

Frith (1989) commented that individuals have a need and desire to achieve high level ‘meaning’. She called this central coherence, and argued that it could be achieved in various ways (e.g., in the form of establishing some gestalt, context, or gist). Essential to her theory is the need to integrate information, which is...
variously described as top-down processing, global processing, parallel processing, processing wholes, and integrating information in context. Her theory is rather general and descriptive in nature which at present is loosely defined because of both a paucity of research and an attempt to offer a broad theory which addresses the strengths and weaknesses in autism.

This capacity for central coherence is proposed by Frith to be diminished in autism, with the result that their processing seems to be devoid of high level meaning. Thus, the central coherence hypothesis makes the prediction that such individuals will experience certain advantages in situations which require them to process in a local or piecemeal way, whilst experiencing certain disadvantages in situations which require them to integrate elements. Several studies support these predictions. Thus, some children and adults with autism show an enhanced ability to detected embedded figures (Jolliffe & Baron-Cohen, 1997; Shah & Frith, 1983), whereas others suggest that children with autism may demonstrate a failure in arguably low level visual integration (Jarrold & Russell, 1997).

In a case report, an adult with autism having normal intelligence experienced difficulties in perceiving the geometric ‘impossibility’ of Impossible Figures (Mottron & Belleville, 1993). Such figures (e.g., the Penrose triangle) are locally congruent, but globally incongruent. Mottron and Belleville believe that this difficulty stems from a defect in integrating parts of the figure, a process upon which the detection of geometric impossibility depends. Furthermore, Mottron (Mottron, Belleville, & Ménard, 1996) found that unlike normally intelligent controls, individuals with autism were not hindered by the three dimensional impossibility of Impossible Figures, when the time taken to copy the figure was used as a performance measure. Control individuals with normal intelligence took significantly longer to draw these figures, whereas they had similar response times when drawing possible figures, objects and non-objects.

Both pieces of evidence suggest that individuals with autism do have difficulty in integrating elements into gestalts. This evidence supports the central coherence theory (Frith, 1989) at the visual perceptual level, which involves low level meaning. There only seems to be one study which may address the central coherence hypothesis at the visual conceptual level (Jarrold & Russell, 1997). This is despite the fact that weak central coherence is thought by Frith to be cognitive in nature and thus involves a dysfunction in central thought processes. This paper seeks to fill a gap in the literature, by investigating whether individuals with autism, or the related condition of Asperger syndrome, have an impaired ability to conceptually integrate visual information. In order to test this, a modified version of the Hooper Visual Organisation Test (HVOT; Hooper, 1983) was used.

The HVOT was originally developed for use with patients with organic brain conditions as a measure of their ability to mentally integrate separate visual elements. Therefore, the test examines visuoconceptual functions. It consists of
30 line drawings depicting simple objects. Each line drawing is cut into pieces and arranged in a puzzle-like fashion. The patient is asked to identify what each object would be if it were put together correctly. The objects used are familiar to most people and would be easily recognised if shown in their usual configuration.

Walker (1956, 1957) and Nadler (Nadler, Grace, White, Butters, & Malloy, 1996) found that whereas brain damaged groups did not differ significantly in terms of the quantitative scoring procedures, they often differed significantly in terms of one of Hooper’s (1983) qualitative scoring categories. Hooper differentiated four major qualitative categories: isolate, perseverative, bizarre, and neologistic responses. ‘Isolate responses’ refers to the use of only one part of the picture to formulate a response, for example, a ‘pipe’ for the tail of a mouse (Walker, 1957). Such a response is said to occur with patients who exhibit fragmentation. ‘Perseverative responses’ are of two kinds: (1) repeating a previously correct or incorrect response on a later item; or (2) providing responses that are unrelated to the current stimulus item but are related to a previous item, for example, ‘pliers’, ‘drill’, ‘screwdriver’, for subsequent items after successful identification of the ‘hammer’ item. Such responses are relatively rare, particularly with adults. ‘Neologistic responses’ are combinations of letters which form a nonsense word, for example, ‘shoon’, for a hammer. Patients who show qualitative anomalies tend to identify most of the items correctly, thus demonstrating that they understand the instruction to integrate fragments. Yet, on one or more of the three items that contain one piece most clearly resembling an object in its own right, patients who have a tendency to view their world in a very fragmented manner will interpret that one piece without attending to any of the others in the item.

Some single case studies in autism have used the HVOT as part of a battery of tests (Mottron & Belleville, 1993; Stevens & Moffitt, 1988). For instance, an adult with Asperger syndrome produced 5 ‘isolate’ errors, which is said to reflect a concrete or fragmented cognitive style and a difficulty integrating visual information holistically (Stevens & Moffitt, 1988). In contrast, another adult with high functioning autism was found to be unimpaired (Mottron & Belleville, 1993). Given that these were single case studies and had conflicting results, interpretation of these results is unclear. Furthermore, the test did not assess central coherence theory directly, and contains some methodological problems; these are addressed next.

The first methodological problem pertains to the nature of the stimuli employed. Of the items that make up this test, half can be recognised from a single element, and therefore do not necessarily require the individual to mentally integrate the elements at all. For instance, the ‘apple’ item can be recognised as such from the top segment, so there is no requirement for this to be integrated with its lower two segments. This is a major drawback for a test which purports to assess visual integration abilities. Second, this test is not a
pure measure of visual integration ability because performance measures are confounded by object recognition abilities. Specifically, there is no control task which assesses the individual’s ability to name the objects employed in the test, thus one cannot be sure whether deficits on this task reflect impairments in visual integration or impairments in object recognition.

The experiment reported in this paper modifies the HVOT in two main ways. Using a completely new set of stimuli, two conditions were created. One condition presented fragments of objects which had to be integrated for the object to be successfully identified. Thus, none of the objects in this condition could be recognised from a single element. This was established through a pilot study. The second condition presented just a single element or part of an object and (again, as established through a pilot study) this object could be identified from the single part presented. A quarter of the stimuli in the first condition (the multiple fragments condition) had an element which was meaningful, which could give rise to an incorrect interpretation of the item if an individual had a tendency to make isolate errors. Inclusion of these type of stimuli was necessary not only to check for problems with fragmentation, but also to check for problems with ‘mental disengagement’, as an individual who has trouble mentally disengaging from a fragment may be impaired at integrating multiple fragments, and particularly so on items where one element was particularly salient (Russell, 1997).

This modified HVOT thus tests two cognitive abilities: integration of elements; and recognition of objects. Integration of elements requires participants to mentally integrate the elements from several locations so as to construct a structural description of the object. Object recognition requires matching this structural description with a stored visual representation, so as to identify the object. Because the single part in the ‘part’ condition does not require the binding of elements to create a structural description (as it is recognisable by itself), it is sufficient in itself to be matched with a stored visual representation, in order to identify what is being portrayed.

If individuals with either autism or Asperger syndrome have weak central coherence, then this theory predicts that they will be less proficient at conceptually integrating the elements so as to construct a structural description. The central coherence theory would not, however, predict a problem in matching this description with a stored visual representation, so object recognition (performance on the control task) should be unimpaired. These predictions were tested here.

**PARTICIPANTS**

Over the last twenty years there has been much debate about whether or not there is a unitary condition of autism that varies in severity levels, or whether there are different types of autism that form part of a spectrum. In the last ten to fifteen
years there has been increasing use of the label ‘Asperger syndrome’, which has as one of its diagnostic requirements no clinically significant delay in early language development (ICD-10; World Health Organisation, 1992). Thus, single words have to be used by two years and phrase speech by three years. It seems then that this type of autism is distinguishable from Kanner’s classical cases all of whom had clinical delays in early language development (Kanner, 1943).

Recently, there has been a lot of interest both in defining autism and in interpreting diagnostic systems. The DSM-IV (American Psychiatric Association, 1994) makes clear that one can have autism with and without language delay based on the number and type of symptoms shown. Thus, depending on the number and type of symptoms shown, a diagnosis of autism can prevail over Asperger syndrome. We chose to distinguish individuals with a history of autistic disorder on the basis of their early language development. This gave rise to two groups one of which clinicians regarded as meeting the criteria for Asperger syndrome and the other which clinicians regarded as meeting the criteria for autism (DSM-IV).

For every participant, clinicians were contacted (with the patients’ and parents’ consent) and medical files were inspected. In cases where either the clinician or the parents could not be sure about early language development these individuals were excluded. To assist in this checking process parents were also given the revised Howlin (1995) screening questionnaire to complete. (This was devised at the Maudsley Hospital, London, makes use of DSM-IV criteria and seeks to identify the presence of autistic symptomatology and whether there was a clinically significant delay in early language development.) There was agreement on early language development in all but one case, and this case was excluded. Individuals whose early language development was considered to be borderline were also excluded.

Inevitably, the Asperger group contained a number of individuals who had not received a diagnosis until adulthood, the remainder were diagnosed in childhood or adolescence. Thus, for the former group, clinicians made their diagnosis retrospectively. For the latter group, and for a couple of individuals, clinicians revised their diagnosis of autism to that of Asperger syndrome on the basis that these individuals did not have a clinically significant delay in early language development. Whereas it is possible that the DSM-IV might regard some of the Asperger participants as having autism without language delay, we have not questioned the clinicians’ diagnosis of Asperger syndrome, especially as this leaves unaffected the key differentiator between the two groups; early language development. Thus, in the autism group all were clinically delayed in language development and would therefore have been considered to have had a history of classical autism. Whereas in the Asperger group none were clinically delayed in their language development.

A total of 51 adults participated in this experiment. These comprised 17 with autism, 17 with Asperger syndrome, and 17 adult control participants with normal
intelligence. The controls acted as a comparison group for the two clinical groups and were taken from the general population of Cambridge.

All participants were required to be of at least normal intelligence (i.e., scoring $\geq 85$) on the WAIS-R (Wechsler, 1981, full-scale, performance and verbal IQ). All three groups prior to their recruitment were screened to check whether they had any history of psychiatric disorder, neurological disorder, or a head injury. Individuals were excluded if they reported any of these factors, and for the clinical groups parents and professionals were consulted. All participants were also required to be medication-free at the time of testing. There was also screening criteria specific to the clinical and control groups. The control group had to be free of any family history of autism or Asperger syndrome. The clinical groups were selected on the basis of their ability to pass both first order and second order belief tasks, and they were screened to exclude those who might be depressed, as depression is much more common in autism and Asperger syndrome and can effect judgement as well as social functioning.

The control participants were chosen to match the clinical groups as closely as possible with respect to the characteristics of age, IQ, sex, and handedness. Because of the difficulty matching on all these measures, participants were not pairwise matched, and rather, a stratified approach was employed. Table 1 gives

<table>
<thead>
<tr>
<th>TABLE 1</th>
<th>Characteristics of participants</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>Group</strong></td>
<td><strong>CA</strong></td>
</tr>
<tr>
<td>Normal ($n = 17$)</td>
<td>30.00</td>
</tr>
<tr>
<td>Autism ($n = 17$)</td>
<td>30.71</td>
</tr>
<tr>
<td>Asperger ($n = 17$)</td>
<td>27.77</td>
</tr>
</tbody>
</table>

CA, chronological age; VIQ, verbal IQ; PIQ, performance IQ; FSIQ, full-scale IQ.

1 Participants were given first and second order theory of mind tests. The first order task was a version of Perner’s (Perner, Frith, Leslie, & Leekam, 1989) Smarties task. The second order task was Baron-Cohen’s (1989a) ice cream van test. Whereas all participants passed the first order task, 5 out of 51 participants failed the second order task. These included 1 participant with Asperger syndrome, 2 with high-functioning autism, and 2 control participants with normal intelligence. These participants were then retested on a new variation of the second order belief task (designed by the first author) and all were found to pass.
the participant details of chronological age (CA), verbal IQ (VIQ), performance IQ (PIQ), and full-scale IQ (FSIQ). Four 1-way ANOVAs revealed no significant differences between groups on any of these variables: CA: $F(2, 48) = 0.59, p = .56$; VIQ: $F(2, 48) = 0.51, p = .60$; PIQ: $F(2, 48) = 0.58, p = .57$; and FSIQ: $F(2, 48) = 0.10, p = .91$.

Over the past two decades it has been assumed that individuals with autism have a verbal IQ that is less than their performance IQ. However, the presumed typical pattern was not found in our sample of high-functioning adults with autism. This failure to find a relatively lower verbal ability is reminiscent of a number of studies that similarly fail to document this pattern (Ehlers et al., 1997; Lockyer & Rutter, 1970; Manjiviona & Prior, 1995; Minshew, Goldstein, Muenz, & Payton, 1992; Rumsey & Hamburger, 1990; Tymchuk, Simmons, & Neafsey, 1977). Of interest is the same finding from a recent study that was set up specifically to test whether VIQ < PIQ in a very large group of rigorously diagnosed high-functioning children and adults with autism (Siegel, Minshew, & Goldstein, 1996). Thus, the presumed prototypic pattern has not been found consistently. Rumsey (1992) suggested that in autism any unique pattern in test scores may depend on ability level, such that VIQ and PIQ differences become smaller or may not be present in samples with average range IQ scores. Rumsey’s point is important because our sample was composed of individuals with average or above average intelligence.

The sex ratio in all three groups was 15:2 (m:f), reflecting the sex ratio found in these clinical groups in other studies (Klin, Volkmar, Sparrow, Cicchetti, & Rourke, 1995; Wing, 1981). The sex ratio was the same across groups, because the Modified Hooper test has a spatial element to it and aspects of spatial ability are known to be superior in males (Maccoby & Jacklin, 1975; Voyer, Voyer & Bryden, 1995). The groups were closely matched on handedness, there being 15 right-handed and 2 left-handed individuals in the normally intelligent control and high-functioning autism group, and there being 14 right-handed and 3 left-handed in the Asperger group. All participants were born in England and English was their first language. All three groups contained participants from various socio-economic backgrounds and the three groups were broadly equivalent in terms of educational attainment (five or six individuals within each group were educated to university level; four or five were educated to A-Level or BTEC standard, four or five to GCSE or O-Level standard; two or three had no formal qualifications). Eight of the individuals in each group were in employment and eight or nine were seeking full- or part-time employment.

The majority of clinical participants were tested in their place of residence, except where some preferred to be tested at the university. All control participants were tested in a quiet room at the university.
MATERIALS

The stimuli presented were black line drawings of objects. Each object appeared on a separate plain white card measuring $9 \times 9$ cm. In the initial piloting phase, the stimuli were presented to nine adults with normal intelligence free from disability or illness, in order to ascertain whether they were relatively easy to name. As the objects were easy to name they were used in the second piloting phase. Half of these objects were presented fragmented into pieces (ranging from 2 to 5 fragments) which had been tilted with respect to each other. The other half presented just a single part of an object. These sets were presented to 20 different control subjects with normal intelligence who had to identify what the objects were from the fragments and a single part. From this second piloting phase items were selected to make both sets comparable in terms of frequency and difficulty, as both sets contained different items. The frequencies were based on norms taken from an English research corpus (TEC, 1999). The difficulty was determined by selecting stimuli that resulted in the performance means and frequency ratings being approximately equal on the fragments and part sets. Thus, items were selected which resulted in the performance means on each set being the same when rounded to the nearest whole number. Similarly, the selected items took into account the frequency of the stimuli such that the mean frequency ratings did not differ between sets: $t(30) = 0.08, p = .93$ two-tailed.

Whilst matching the sets on mean frequency ratings the stimuli within each set had differing frequencies. This reflects the natural world where objects do not occur with equal frequency, and it also avoided the test being rendered too easy or too difficult as a result of pitching frequencies at a certain level. Examples of the stimuli used in this experiment are shown in the Appendix.

The test stimuli consisted of 32 of the original objects. The 32 objects constituted two sets of 16 objects in each. The objects in each condition were different but had been matched on frequency and difficulty. Also these sets consisted of the fragmented objects (F) and a single part of an object (P). Again the objects in their complete form were presented, but this time as naming controls for the F and P sets—the naming control sets being labelled FNC and PNC, respectively. Thus the test stimuli consisted of four sets of test cards: the F and P sets and the FNC and PNC sets. Three trial items for each set were employed as training stimuli for the test items.

A stopwatch was used to determine the length of time it took the participant to determine what each object was. In order to assist in the recording of response times, there was a large square of blank white card, the purpose of which was to cover the test cards. This could be removed rapidly, and therefore allowed greater accuracy in recording response times than could be achieved if one just turned over the test card and then starting the stopwatch.
PROCEDURE

Each participant was tested individually in a room free from distractions. The experimenter sat opposite the participant, so that she could conceal more easily the names of the objects which were to be checked off on a score sheet. Four sets of cards were presented, one set at a time. The set being presented was placed face up on the table directly in front of the person, and on top of this pile of cards there was the large square of card, which was used to control participants’ response times with greater accuracy.

The trial and test items

The three trial items of the Fragments set were presented first and in a fixed order to each participant. The participant was told that he/she was going to see some pictures of objects and that they were to look at some trial items before beginning the actual test items. The first item to be uncovered was a balloon which had been fragmented into two pieces. Participants were asked what the object would be if the two pieces were put together. The participant was then shown this item made up from the trial items of set FNC. The experimenter then presented in an identical manner the remaining two trial items. All participants found the trial items relatively easy.

After the three trial items, the experimenter informed the participant that the procedure for the test items would be the same as for the trial items, except that now their responses would be timed and that they would not see the items complete until the end of the test when they would be asked to name them. They were also informed that they could not amend their decision as to what the object was, otherwise their response times would be distorted. Participants were then given the shuffled Fragments set. Each participant received a different random order in order to ensure that there would be no order effects. Furthermore, after each response to an item, the loose card which covered the pile of cards was replaced back on top of the pile and the top card just viewed was slid away from underneath and placed to one side face downwards on the table. This made it possible for the next card to be ready for presentation.

On completion of the Fragments set, participants were given the trial items of the Part set. These were introduced in a similar manner to the previous set’s trial items, but here, participants were told that with these cards they would now be seeing just a part of an object and that they were to guess what the whole object was just from seeing this part. After completing the three trial items of this set without experiencing any difficulty, participants were given the test items.

Each participant always received the Fragments set, followed by the Part set, as there was a strong possibility that had the clinical groups been given the Part set first, they could get used to activating a particular response schema (i.e., that of attending very closely to a single element). Such a response bias could
prevent the visual integration of multiple fragments which was the requirement of the fragment integration task. Because it was predicted that the clinical groups might be less efficient on the fragment integration condition, having this condition first and the ‘part’ condition second enabled the latter to act as a control for both attentional or fatigue effects.

Participants were finally given the naming control sets. For consistency, the Fragments Naming Control Set was always presented first, followed by the Part Naming Control Set, and the pictures within each set were again randomised for each participant. Participants were told that they had to simply name the objects portrayed. This easy task functioned solely to ensure that any apparent deficit in visually integrating fragments was not in fact due to either unfamiliarity with some of the objects used in the test, or to a deficit in naming whole object representations per se. Participants’ responses were not timed on these sets, as it was only important to assess naming.

Scoring

The recording of response times was achieved as follows: Timing was started as soon as the blank card was removed from the pile of test cards. When the participant reported what the object on the top of the pile was, the experimenter recorded the response time. If the participant made an error, or omitted to make a response, this was also recorded. All response times were scored in seconds and hundredths of a second, as participants could recognise some of the objects very rapidly. It was assumed that any timing inaccuracies resulting from the experimenter’s operating of the stopwatch would average out across the test items. Participants were given as long as they wanted to make their decision. The timing was to ensure that the clinical groups were not failing due to being impulsive, which would be expected to result in faster response times overall, and particularly so on failed items. Responses that were synonyms (e.g., ‘settee’ and ‘sofa’ were classed as correct).

RESULTS

Accuracy

Accuracy scores were the number of correctly recognised objects, which could range from 0 to 16 for each of the two conditions (see Table 2).

The accuracy scores were approximately normally distributed and the variances were approximately equal, so a 2-factor repeated-measures ANOVA was performed on the mean accuracy scores for each of the three groups. This ANOVA had a between-participant variable of Group, and a within-participant variable of Condition (Fragments and Part). The ANOVA revealed significant main effects of Condition: $F(1, 48) = 30.48, p < .01$, and Group: $F(2, 48) = 6.04,$
There was also a higher order interaction of Group by Condition, which was as predicted significant: $F(2, 48) = 19.75, p < .01$, (see Figure 1). This Group by Condition interaction means that this interaction must at least in part be due to there being different Group effects for the two conditions. To examine whether the general group effect applies to one or more of the conditions, simple effects were examined which compared the different Groups on each Condition. As predicted, analysis of simple effects showed the effect of Group to be significant only for the Fragments condition: $F_{\text{Fragments}} (2, 48) = 16.19, p < .01$; $F_{\text{Part}} (2, 48) = 0.24, p = .79$. The Group effect and this Group by Condition interaction was investigated further using t-tests. Planned contrasts of the cell means indicated that the Fragments mean of the autism and Asperger groups were as predicted significantly lower than that of the control group with normal intelligence: $t_{\text{aut.}} (48) = 5.60, p < .01$; $t_{\text{Asp.}} (48) = 2.84, p < .01$. We also found an unpredicted difference on the Fragments mean between the autism and Asperger groups: $t(48) = 2.76, p < .05$, which suggests that the autism group was significantly less efficient than the Asperger group.

Given that the Condition effect was significant, it is useful to examine whether there were different Condition effects for the three Groups. Simple effects were examined which compared the two conditions for each group. Analysis of simple effects showed the effect of Condition to be significant for the autism group: $F(1, 48) = 56.46, p < .01$, and the Asperger group: $F(1, 48) = 11.67, p < .01$, but not the control group with normal intelligence $F(1, 48) = 1.87, p = .18$. Observation of the means (see Table 2 and Figure 1) show the autism and Asperger groups to be significantly worse on the Fragments condition relative to their own performance on the Part condition, whereas the control group were more level in their performance on both conditions, being neither significantly better nor significantly worse on either of the two conditions.

### Table 2

<table>
<thead>
<tr>
<th>Group</th>
<th>Fragment</th>
<th>Part</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal</td>
<td>Mean 12.06</td>
<td>11.59</td>
</tr>
<tr>
<td>(n = 17)</td>
<td>SD 1.25</td>
<td>1.54</td>
</tr>
<tr>
<td></td>
<td>Range (9–14)</td>
<td>(9–14)</td>
</tr>
<tr>
<td>Autism</td>
<td>Mean 8.59</td>
<td>11.18</td>
</tr>
<tr>
<td>(n = 17)</td>
<td>SD 2.29</td>
<td>1.74</td>
</tr>
<tr>
<td></td>
<td>Range (5–12)</td>
<td>(8–14)</td>
</tr>
<tr>
<td>Asperger</td>
<td>Mean 10.12</td>
<td>11.29</td>
</tr>
<tr>
<td>(n = 17)</td>
<td>SD 1.58</td>
<td>2.08</td>
</tr>
<tr>
<td></td>
<td>Range (8–12)</td>
<td>(7–14)</td>
</tr>
</tbody>
</table>
Figure 1. Effect of condition on mean accuracy scores.
It is important to determine whether the clinical groups’ inefficiency in integrating fragments is determined by just a few individuals in each group or characterised the majority of these groups. To this end, the number of participants in each group scoring above (and below) the control group mean was calculated. This was compared to the numbers of participants in the control group scoring above (and below) their mean. The analysis revealed that the autism and Asperger groups differed significantly from the control group [Yates Continuity Correction to correct for expected frequencies < 5, \( \chi^2(1) = 6.48, p < .05 \) for both].

Response time

Because it was possible that the clinical groups could fail the fragments (and part) set due to being impulsive, response times for correct and incorrect responses were examined (see Table 3). Examining first the fragmented set: The data were approximately normally distributed for each group and the variances were approximately equal, so two 1-factor ANOVAs were performed, one on each type of response. These ANOVAs revealed that the groups did not differ in response time on the number of correct or incorrect responses: \( F_{\text{correct}}(2, 48) = 0.96, p = .39; F_{\text{incorrect}}(2, 48) = 0.06, p = .95 \). Examining the Part set: The data were again approximately normally distributed for each group and the variances were roughly equal, so two 1-factor ANOVAs were performed, one on each type of response. These ANOVAs again revealed that the groups did not differ in response time on the number of correct or incorrect responses: \( F_{\text{correct}}(2, 48) = 0.69, p = .51; F_{\text{incorrect}}(2, 48) = 1.38, p = .26 \). Whether one looks at the mean response time on the correct or incorrect responses, the analysis yielded the same results. The clinical groups did not differ in their response times.

<table>
<thead>
<tr>
<th>Group</th>
<th>Fragment Incorrect</th>
<th>Fragment Correct</th>
<th>Part Incorrect</th>
<th>Part Correct</th>
</tr>
</thead>
<tbody>
<tr>
<td>Normal</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(n = 17) Mean</td>
<td>16.97</td>
<td>4.22</td>
<td>12.83</td>
<td>2.95</td>
</tr>
<tr>
<td></td>
<td>SD</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>7.51</td>
<td>1.84</td>
<td>6.24</td>
<td>1.24</td>
</tr>
<tr>
<td></td>
<td>Range</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(6.21–30.12)</td>
<td>(2.32–8.93)</td>
<td>(3.14–26.38)</td>
<td>(1.05–5.02)</td>
</tr>
<tr>
<td>Autism</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(n = 17) Mean</td>
<td>16.17</td>
<td>5.16</td>
<td>9.54</td>
<td>3.79</td>
</tr>
<tr>
<td></td>
<td>SD</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>9.89</td>
<td>3.23</td>
<td>6.18</td>
<td>3.52</td>
</tr>
<tr>
<td></td>
<td>Range</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(5.15–41.19)</td>
<td>(2.09–15.00)</td>
<td>(3.33–24.14)</td>
<td>(1.32–16.70)</td>
</tr>
<tr>
<td>Asperger</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>(n = 17) Mean</td>
<td>17.05</td>
<td>5.42</td>
<td>11.22</td>
<td>3.81</td>
</tr>
<tr>
<td></td>
<td>SD</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>7.84</td>
<td>2.68</td>
<td>4.81</td>
<td>1.98</td>
</tr>
<tr>
<td></td>
<td>Range</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td>(6.34–33.30)</td>
<td>(1.59–13.00)</td>
<td>(4.29–19.99)</td>
<td>(1.25–8.59)</td>
</tr>
</tbody>
</table>
Object naming, omissions, and disengagement errors

Because it is the Fragments condition which is so difficult for the clinical groups the remaining analysis will focus on this condition. Because the clinical groups’ relative difficulty with the Fragments condition could simply be because they had difficulty naming some of the objects portrayed, the naming errors were examined. The number of naming errors on the Fragment Naming Control set (FNC) were calculated for each group and contrasted with each group’s total number of errors. A $3 \times 2$ Pearson chi-square test was then conducted to see whether the groups differed in the proportion of their errors which could be considered to be due to an inability to name (and hence know) the objects portrayed. Chi-squared analysis revealed no significant difference between the three groups: $\chi^2(2) = 0.03$, $p = .98$.

It was also possible that the clinical groups’ relative difficulty with the Fragments condition was simply because they were less willing to give a response or make a guess in comparison to their control group (i.e., they were tending to make more omissions for whatever reason), rather than there really being a problem in integrating fragments. This was thought possible given their communication difficulties, lack of spontaneity, and their problems with generating ideas, so the number of omissions were examined. The number of omissions made by each group for the Fragments set were calculated and contrasted as a proportion of their total number of errors. A $3 \times 2$ Pearson chi-square test was again conducted to see whether the groups differed in the proportion of their errors which could be considered to be due to omissions. Chi-squared analysis revealed no significant difference between the three participant groups: $\chi^2(2) = 0.04$, $p = .98$.

Finally, it is possible that the clinical groups’ relative difficulty with the Fragments condition might be due to a problem in mentally disengaging from a fragment. A quarter of the stimuli had an element which in itself was meaningful, so a tendency to be captured by the visual salience of an element would be expected to show up most on these items. Thus, the row of keys of the accordion could be described as a piano, the head of the unicorn (minus his horn) as a horse, the bottom of the rake as a comb, the body of the jug as a vase, and the base of the frying pan as a bowl. The actual number of disengagement errors were quite small and did seem to occur only with these stimuli. Nevertheless, these errors were calculated for each group and contrasted as a proportion of their total number of errors. A final $3 \times 2$ Pearson chi-square test was conducted to see whether the groups differed in the proportion of their errors which could be considered to be due to a problem with mental disengagement. Chi-squared analysis revealed no significant difference between the three participant groups: $\chi^2(2) = 0.09$, $p = .96$. 
As predicted by central coherence theory (Frith, 1989), both groups of clinical participants (with either autism or Asperger syndrome) were significantly impaired in their ability to integrate fragments holistically to identify an object. However, they were unimpaired in their ability to recognise an object from a single part. The impairment in integration replicates the finding of Stevens and Moffit (1988) who found an adult with Asperger syndrome to have a borderline impairment. This is at odds with the study of Mottron and Belleville (1993) who found no such impairment in an adult with high-functioning autism. The differences between these earlier studies and the modified version of the HVOT probably lies in the methodological differences between the two tasks employed.

The modified Hooper test sought to directly test central coherence theory, whereas the earlier studies did not. This led to the adoption of two conditions, one of which required the integration of multiple fragments (Fragments condition), and the other of which required only a consideration of a single element or part (Part condition). Individuals with either autism or Asperger syndrome showed evidence of relatively impaired processing on the Fragments condition, but normal processing on the Part condition. This suggests that the failure of the Mottron and Belleville study to detect a difference, and the mild impairment found in the Stevens and Moffit study, was confounded by the fact that half of the stimuli they used could be recognised from a single part, and therefore resembled our Part condition.

The visuoconceptual deficit observed in the clinical groups raises the question as to how this task differs from that of the Object Assembly task in the Wechsler (1974, 1981) IQ tests, as they both have puzzle-like pieces, and in both, a whole needs to be constructed from parts. Individuals with autism are not usually found to have deficits on Object Assembly (Happé, 1994), and are known to be proficient at jigsaws (Baron-Cohen & Bolton, 1993). One reason may be that it is possible to construct the design in the Object Assembly task in a ‘bottom-up’ way, on the basis of local connections (contours) between the puzzle pieces. It is not until the object is nearly completed that the identity of the most difficult items becomes apparent (Lezak, 1995). Thus, individuals who are unable to conceptualise what the Object Assembly constructions should be, could put them together in a piecemeal fashion, by matching lines and edges in a serial manner. The Hooper test does not afford this opportunity. Indeed, for many of the items it is not possible to see where the pieces would join if one were to try to use such a serial approach. The Hooper test requires the fragments to be integrated in parallel (because the lines do not facilitate an obvious local matching procedure), a process which is essentially ‘top-down’ rather than bottom-up. It is for this reason that the original Hooper test was described as providing a quantitative measure of visuoconceptual functioning.
Further evidence for the top-down (and hence parallel processing) nature of the task comes from response times. Participants from all groups tended to identify many of the items very rapidly. This suggests processing was in parallel and hence top-down. However, inspection of the clinical groups’ data revealed they had a tendency to take longer as the number of elements to be integrated increased. This might suggest their difficulty in integrating multiple elements, and that where they failed to employ a top-down or parallel processing strategy, they had to resort to using a serial matching approach, as one would use when physically matching puzzle pieces. Equally striking is the finding that although the clinical groups tended to respond more slowly on the items that had multiple elements, and the items that they could not identify, they were not across all stimuli significantly slower than their matched control group. This would seem to go against the predictions of weak central coherence, as a difficulty integrating elements would predict that they should be slower. This seems unlikely to be due to impulsivity or poor motivation for the reasons explained a little later, however, it might be due to the fact that there were too few items with multiple elements which would enable us to successfully demonstrate this prediction. This could be clarified further with additional research which contrasts response times for items with multiple elements, with that of items with just a couple of elements. Also of interest is the finding that despite the clinical groups’ difficulty integrating elements they were not observed to be less accurate as the number of elements to be integrated increased. This also is quite hard to explain, as, again, weak central coherence would suggest that their performance should be less accurate as the number of elements to be integrated increased. Again, this may be due to the fact that there were too few items with multiple elements. This clearly is another area for future study.

When looking at the performance within groups it was found that the clinical participants performed significantly worse on the Fragments relative to their own performance on the Part condition. On the other hand, the control group performed similarly on both conditions. It could be interpreted that differences in central coherence ability underlies the performance of the clinical groups and the control group with normal intelligence. Thus, the control group’s strong coherence enabled them to cope with the demands of conceptually integrating fragments, whereas the clinical groups’ weak central coherence prevented them from performing as well on the Fragments condition as they did on the Part condition. In this regard it is interesting that the autism group is significantly less able to integrate elements than the Asperger group. This would be consistent with the anecdotal reports of individuals with autism.

In the second part of this discussion we turn to consider other factors which might be influencing the clinical groups’ ability to integrate elements. The first of these pertains to motivational problems. It is unlikely that the clinical groups’ poor performance is a secondary consequence of motivational factors. Three separate results argue against this hypothesis. First, they were not impaired on
the control or Part task and one would expect motivational differences to also be evident on this task. Second, participants with autism or Asperger syndrome were able to integrate elements together so as to identify some of the more difficult to identify objects, thus displaying motivated behaviour. Third, the finding most pertinent to a motivational hypothesis is that (despite their poor overall performance) the clinical groups were observed to respond to the varying degrees of difficulty of the stimulus materials employed in much the same way as their control group. In fact, even the items that they failed, demonstrated their highly motivated behaviour as reflected in the good (albeit incorrect) responses they made. For example, some of the items that were incorrect but good guesses were as follows: a paper bag for the sweet, an axe for the broom, a chair or slide for the athletics’ jumping hurdle, and a rod or musical instrument for the walking stick. Such evidence leads one to conclude that the clinical groups’ difficulty in mentally synthesising fragments in the Fragments condition cannot be due to motivation problems. Such findings also rule out inattention and, in particular, fatigue playing a role as the control or part task was always presented after the critical or fragments task, so fatigue effects would be expected to be more evident on this task.

A second factor which could be suggested as playing a contributory role in the clinical groups’ deficit in integrating elements is impulsivity. However, it is unlikely that clinical groups perform poorly simply because they give the first response which comes into their head, as like the control group, they tended to respond rapidly on the objects that they could identify easily, but tended to respond more slowly on the items they could not identify easily. Furthermore, like the control group they took longer to make their response on the Fragments condition than they did on the Part condition. This was presumably because of the increase in the number of elements which needed to be considered and because of their difficulty with this condition. The clinical groups’ response times on the Fragments condition does not support impulsive responding.

A third factor which might have given rise to a problem in integrating multiple elements stems from the anecdotal reports which suggest that children with autism fixate on small morphological details (Rimland, 1971) together with the experimental evidence that suggests they have problems with mental disengagement (Hughes & Russell, 1993; Russell, Mauthner, Sharpe, & Tidswell, 1991). However, there are number of studies that demonstrate that children with autism can mentally disengage. These include visual perspective-taking tasks, where the child has to infer what someone else can see from his/her spatial position, even if this is different to what the child currently sees (Baron-Cohen, 1989b; Hobson, 1984; Tan & Harris, 1991). Similarly, they seem to perform at a level appropriate to their mental age on judging the content of ‘false’ photographs, maps, drawings, and models, all of which require the child to put aside his/her knowledge of current reality in order to answer a question about an out-of-date situation (Charman & Baron-Cohen, 1992, 1995; Leekam
Perner, 1991; Leslie & Thaiss, 1992). The modified Hooper task itself also contains a quarter of the fragmented stimuli with an element which in itself was meaningful, so a tendency to be captured by the visual salience of an element (in this case a meaningful element), or to make an isolate response, would be expected to show up most on these items. However, the number of disengagement or isolate errors did not differ between groups. The individual with Asperger syndrome in the Stevens and Moffitt study gave five isolate errors (out of a total of 30 stimuli). This may have something to do with the fact that about half of their items could be identified from a single part, which would encourage such a response bias. This did not occur in the study presented here, possibly because the fragments and single part responses were kept separate and the Part condition always followed the Fragments condition. Furthermore, the clinical groups’ attempts to identify the item portrayed, amidst their difficulties with integration, led them to occasionally (albeit rarely) give a meaningful, but incorrect interpretation for each element rather than tending to focus on a single element (e.g., the frying pan became a knife and bowl). On the face of such evidence one is left to conclude that a problem with mental disengagement or fragmentation per se cannot be at the heart of the clinical groups’ difficulty in mentally integrating fragments.

A fourth factor which could account for the problem in identifying the objects in the Fragments condition, is that the clinical participants could have been unfamiliar with the objects employed, or have a problem with object recognition in general. Neither of these factors seem likely for two reasons. First, the participants did not differ in their ability to correctly identify the fragmented objects in their whole form when presented in the Fragments naming control task. Second, their ability to correctly identify the whole objects employed in the Part condition suggests that they do not have a problem with object recognition in general, which is consistent with the evidence for unimpaired object recognition demonstrated in a high-functioning adult with autism (Mottron & Belleville, 1993).

The results strongly suggest that it is the need to integrate the fragments that accounts for the clinical groups’ difficulties. Future work will need to investigate the neural basis for this problem. Recent evidence suggests that patients with right hemisphere lesions are impaired at integrating fragments on the traditional HVOT (Fitz, Conrad, Hom, Sarff, & Majovski, 1992; Nadler et al., 1996), and it is possible that right hemisphere lesions not only cause the

---

2 In this experiment, ‘isolate errors’ as a result of fragmentation problems are assumed to be synonymous with 'mental disengagement errors’. The test does not distinguish between the two and it is not entirely clear what the relationship between the two is, as a difficulty with mental disengagement might be made worse by problems with fragmentation. This in no way creates a problem for this task, since the analysis sought only to demonstrate that ‘integration errors’ did not occur because of a tendency to concentrate on a single element or detail.
integration deficits but also the social and pragmatic deficits characteristic of autism (Brownell, Potter, Bihrle, & Gardner, 1986; Brownell, Potter, & Michelow, 1984; Brownell, Simpson, Bihrle, Potter, & Gardner, 1990; Happé & Frith, 1996).

One of the problems with both this modified and the traditional HVOT is that the stimuli are line drawings, and are not, therefore, three-dimensional objects in the real world. It is possible that the integration of fragmented ‘real’ objects might be different. A further consideration for future research, will be to test and contrast the ability to conceptualise the fragments of a real object, with the ability to physically put the object together. Also, future research should correlate performance on the modified HVOT with symptom severity in an attempt to explain in a more direct fashion how the symptoms of autism and Asperger syndrome might be linked to weak central coherence.

Although both groups demonstrated an impairment, the individuals with autism were significantly more impaired than those with Asperger syndrome. In addition, there was wide variability within as well as between clinical groups, which suggests that integration (central coherence) ability may vary widely. Although the ability to integrate elements also varied quite widely in the control group (though less than in the clinical groups), the findings reported here do not suggest that those with an autism spectrum disorder are at one end of the normal continuum on this test of visual synthesis. The clinical participants seemed to possess a real impairment, as only a minority of their scores overlapped with control individuals. However, despite this impairment, the performance of our high-functioning adults with autism or Asperger syndrome was one of inefficiency rather than inability. They were not unable to integrate visual elements, just significantly less proficient at doing so. Therefore our clinical participants differ from patients with apperceptive visual agnosa, who cannot synthesise what they see. Such patients indicate awareness of discrete parts of a word or a phrase, or recognise elements of an object, without organising the discrete percepts into a perceptual whole.

The results of the majority analysis suggest that this lack of proficiency in integrating elements applied to the majority of our subjects with autism or Asperger syndrome. Therefore, it seems that difficulties in visual synthesis characterise at least some individuals on the autism spectrum. The performance of the clinical groups supports Frith’s (1989) proposal that central coherence or the ability to integrate information is impaired in autism and Asperger syndrome. Moreover, the findings are in line with evidence from other studies demonstrating that individuals with autism or Asperger syndrome have difficulty in integrating multiple objects (Carpentieri & Morgan, 1994; Jolliffe & Baron-Cohen, 2001; Rumsey & Hamburger, 1988, 1990). The evidence from these studies and the findings reported in this paper suggest that a problem in integrating information characterises individuals on the autism spectrum. Given that the impairment characterised the majority of individuals with autism, the
findings support Frith’s suggestion that a problem in integrating information might be a core deficit in autism.

Manuscript received 26 June 2000
Revised manuscript received 15 December 2000

REFERENCES


APPENDIX

Fragments condition

Fragment naming control

Fragments condition

Fragment naming control
Part condition  Part naming control

Part condition  Part naming control